Infiltration Estimates from Borehole Water Content Data using the Markov Chain Monte Carlo Method

Zhiming Lu and Bruce A. Robinson

Hydrology, Geochemistry, and Geology Group
Los Alamos National Laboratory

Abstract

A recently developed inversion technique based on the Markov Chain Monte Carlo (MCMC) method is applied to the interpretation of water content data from boreholes around the Laboratory for the purpose of obtaining estimates of the local infiltration rate. The MCMC method provides better estimates of infiltration rate and its uncertainty than existing inverse modeling methods, principally because it takes into account heterogeneity in the medium properties. Numerical experiments with the method illustrate that including small-scale heterogeneity is important for improved matches to the data. When applied to the data from wells in the vicinity of LANL, the method provides reasonable estimates of uncertainty for wells in a variety of topographic settings. Infiltration rates obtained from modeling of water content data from wells in Los Alamos and Mortandad Canyons are in general agreement with previous estimates, but in addition, definitive estimates of the uncertainty have been obtained as well. The analyses presented here use a simplified version of the inversion model. The more advanced versions of the method that include transient flow and solute transport should be useful for interpreting data currently being collected as part of the ER activities in Mortandad Canyon and other locations around the Laboratory.

November 9, 2005

Submitted to LANL Groundwater Protection Program to fulfill FY05 Vadose Zone Project deliverable

I. Model Development

In this section, we describe the mathematical model and the inversion model requirements for obtaining estimates of the local infiltration rate in the vicinity of a well in which measurements of water content and chemical concentrations have been made. This theoretical development is then used later to interpret water-content profiles in several wells on the Pajarito Plateau.

I-1. Statement of the Problem

We consider transient flow in variably saturated porous media satisfying the following continuity equation and Darcy's law:

$$-\nabla \cdot \mathbf{q}(\mathbf{x},t) + g(\mathbf{x},t) = C_s(\psi)\frac{\partial \psi(\mathbf{x},t)}{\partial t}, \tag{1}$$

$$\mathbf{q}(\mathbf{x},t) = -K[\psi]\nabla[\psi(\mathbf{x},t) + x_1], \tag{2}$$

subject to initial and boundary conditions

$$\psi(\mathbf{x},0) = \Psi_0(\mathbf{x}), \qquad \mathbf{x} \in \Omega, \tag{3}$$

$$\psi(\mathbf{x},t) = \Psi(\mathbf{x},t), \qquad \mathbf{x} \in \Gamma_D \tag{4}$$

$$\mathbf{q}(\mathbf{x},t) \cdot \mathbf{n}(\mathbf{x}) = Q(\mathbf{x},t), \qquad \mathbf{x} \in \Gamma_N, \tag{5}$$

where $\boldsymbol{q}$ is the specific discharge (flux), $\psi(\boldsymbol{x},t) + x_1$ is the total head, ψ is the pressure head, $\Psi_0(\boldsymbol{x})$ is the initial pressure head in the domain Ω, $\Psi(\boldsymbol{x},t)$ is the prescribed head on Dirichlet boundary segments Γ_D, $Q(\boldsymbol{x},t)$ is the prescribed flux across Neumann boundary segments Γ_N, $\boldsymbol{n}(\boldsymbol{x})=(n_1,\cdots, n_d)^T$ is an outward unit vector normal to the boundary, $C[\psi]= d\theta/\mathrm{d}\psi$ is the specific moisture capacity, θ is the volumetric water content, and $K[\psi]$ is the unsaturated hydraulic conductivity (assumed to be isotropic locally). Both C and K are functions of pressure head and soil properties at $\boldsymbol{x}$. For convenience, they will be written as $C(\boldsymbol{x},t)$ and $K(\boldsymbol{x},t)$ in the sequel. The elevation x_1 is directed vertically upward. In these coordinates, recharge has a negative sign.

It is clear that models are needed to describe the constitutive relationships of K versus ψ and θ_e versus ψ when the flow is unsaturated. No universal models are available for the constitutive relationships. Instead, empirical models are typically employed, including the Gardner-Russo model [Gardner, 1958; Russo, 1988], the Brooks-Corey model [Brooks and Corey, 1964], and the van Genuchten-Mualem model [van Genuchten, 1980]. Most analytical solutions of the deterministic unsaturated flow equations and most previous stochastic analyses used the Gardner-Russo model because of its simplicity. However, it is generally accepted that the more complex van Genuchten-Mualem and Brooks-Corey models may be more flexible than the simple Gardner-Russo model in describing measured data of $K(\psi)$ and $\theta_e(\psi)$. In this study, we use the van Genuchten-Mualem model:

$$K(\boldsymbol{x},t) = K_s(\boldsymbol{x})\sqrt{S(\boldsymbol{x},t)}\{1-[1-S^{1/m}(\boldsymbol{x},t)]^m\}^2 \tag{7}$$

$$S(\boldsymbol{x},t) = \{1+[-\alpha(\boldsymbol{x})\psi(\boldsymbol{x},t)]^n\}^{-m}, \tag{8}$$

where $\psi \leq 0$. In the above, $S(\mathbf{x},t)= \theta_e/(\theta_s-\theta_r)$ is the effective saturation, θ_r is the residual (irreducible) water content, θ_s is the saturated water content, α and n are fitting parameters, and $m = 1 - 1/n$. With (8), $C_s(\mathbf{x},t) = d\theta_e/\mathrm{d}\psi$ can be expressed explicitly as

$$C_s(\mathbf{x},t) = \alpha(\mathbf{x})[n(\mathbf{x})-1](\theta_s - \theta_r)S^{1/m}(\mathbf{x},t)[1-S^{1/m}(\mathbf{x},t)]^m \qquad (9)$$

In this study, θ_s and θ_r are assumed to be deterministic, as their variabilities are likely to be small compared to that of the effective water content θ_e [*Russo and Bouton,* 1992], while the saturated hydraulic conductivity K_s, the pore size distribution parameter α, and the fitting parameter n are treated as random functions. We assume that the log-transformed saturated hydraulic conductivity $f(\boldsymbol{x}) = \ln K_s(\boldsymbol{x})$, the log-transformed pore size distribution parameter $\beta(\boldsymbol{x}) = \ln \alpha(\boldsymbol{x})$, and $\mu(\boldsymbol{x}) = \ln [\mathrm{n}(\boldsymbol{x})-1]$ follow the normal distribution. Now suppose that there are m_f direct measurements on the log hydraulic conductivity, f_i, $i = \overline{1, m_f}$, m_β measurements on the pore-size distribution parameter β, β_i, $i = \overline{1, m_\beta}$, and m_μ measurements on the fitting parameter, μ_i, $i = \overline{1, m_\mu}$. The sampling locations for three different kinds of direct measurements may be different.

The transport of a nonreactive solute can be described by the classical convection-dispersion transport equation [Bear, 1972]:

$$\nabla\left[D_{ij}\cdot\nabla C(\mathbf{x},t)\right] - v\cdot\nabla C(\mathbf{x},t) = \theta\frac{\partial C(\mathbf{x},t)}{\partial t}\ , \qquad (10)$$

subject to appropriate initial and boundary conditions. Here C is the nonreactive solute concentration, D_{ij} is the dispersion coefficient tensor, and $\boldsymbol{v}$ is the seepage velocity, which can be computed from specific discharge $\boldsymbol{q}$.

An alternative way to characterize nonreactive transport is to record the position of a particle at time t that originates from position $\mathbf{a}$ at time $t = t_0$ and is described by the following kinetic equation

$$\frac{d\mathbf{X}(t;\mathbf{a})}{dt} = \mathbf{V}(\mathbf{X}), \qquad (11)$$

with the initial condition of $\mathbf{X}(t_0;\mathbf{a}) = \mathbf{a}$, where $\boldsymbol{V}$ is Lagrangian velocity. Now we are interested in the travel time, the time taken for a particle to travel from $\mathbf{a}$ to a well or to across a control plane that is perpendicular to the mean flow direction and located at some distance from the source. The travel time $\tau = t - t_0$ can be determined from (11) using the particle tracking technique. If m_p particles are released in the flow domain and m_t travel times are recorded for each particle, we could have $m_p \times m_t$ measurements on travel time τ_{ij}, $i = \overline{1, m_p}$, $j = \overline{1, m_t}$.

In this development, in addition to direct measurements on soil properties, it is assumed that we also have m_ψ measurements on the pressure head, m_θ water content measurements, and m_C concentration measurements or $m_p \times m_t$ measurements on travel time τ_{ij}.

I-2. Representation of Soil Properties

As mentioned above, the soil properties p, where $p = f$, β, or μ, are treated as spatially stationary random space functions with mean $\langle p \rangle$, and covariance function $C_p(\mathbf{x},\mathbf{y})$. These properties are represented by n_p basis kernel functions $b_p(\mathrm{x},\chi)$ centered at some fixed spatial locations $\chi_j^{(p)}, j = \overline{1,n_p}$

$$p(\mathbf{x}) = \sum_{j=1}^{n_p} \gamma_j^{(p)} b_p(\mathbf{x}, \boldsymbol{\chi}_j^{(p)}) \tag{12}$$

where $\gamma_j^{(p)}, j = \overline{1,n_p}$, are coefficients to be determined in the inverse procedure. The kernel functions can be chosen, for example, as an exponential function $b_p(\mathbf{x},\boldsymbol{\chi}) = \exp\left(-\sum_{i=1}^{d} | x_i - \chi_i | / \lambda_i\right)$, where d is the number of space dimensions and λ_i is a parameter that controls the influence of the kernels $b_p(\mathbf{x},\boldsymbol{\chi})$ in the i^{th} dimension. For the fixed basis kernel functions, the parameter field $p(\mathbf{x})$ can be computed from $\gamma_j^{(p)}, j = \overline{1,n_p}$.

I-3. Bayesian Inference

Determining the infiltration rate best estimate and uncertainty requires an inverse model algorithm. In this study, a Bayesian approach is taken. The essence of this approach is embodied in Bayes' Theorem, which can be understood as a mathematical description of the learning process. Bayesian statistical inference requires an additional input not needed by frequentist procedures such as maximum likelihood: *a prior* probability distribution for parameter γ, which embodies our judgment before seeing any data D of how plausible it is that the parameters could have values in the various regions of parameter space. The introduction of a prior is the crucial element that converts statistical inference into an application of probabilistic inference. When we combine a prior distribution $\pi(\gamma)$ for the parameters with the conditional distribution for the observed data, we get a joint distribution for all quantities related to the problem:

$$\pi(\gamma, D) = \pi(\gamma)\pi(D \mid \gamma) = \pi(D)\pi(\gamma \mid D) \tag{13}$$

From this equation, we can derive Bayes' rule for the posterior distribution of the parameters given observed data D:

$$\pi(\gamma \mid D) \propto L(D \mid \gamma)\pi(\gamma) \tag{14}$$

where $L(D|\gamma)$ is the likelihood function. For our problem described above, the likelihood function $L(f_o,\beta_o,\ \mu_o;\ \psi_o,\theta_o,\ \tau_o \mid \gamma)$ of observed data $D = (f_o,\beta_o,\ \mu_o;\ \psi_o,\theta_o,\ \tau_o)$ given parameters γ may be written as

$$L(D \mid \gamma) \propto \exp\left\{-\frac{1}{2}\sum_{p=f,\beta,\mu}\left(p_o - p_m\right)^T \sum\nolimits_p^{-1}\left(p_o - p_m\right) - \frac{1}{2}\sum_{h=\psi,\theta,\tau,C}\left(h_o - h_m\right)^T \sum\nolimits_h^{-1}\left(h_o - h_m\right)\right\} \tag{15}$$

where $\sum_p^{-1}$ is an $m_p \times m_p$ matrix determined by observation errors and representativeness of measurements for parameters $p = f$, β, and μ, $\sum_h^{-1}$ is an $m_h \times m_h$ matrix accounting for

observation error and model discrepancy on dependent variables h = ψ, θ, τ, or C. For the Bayesian approach, we need to specify a prior distribution for γ. One such example is

$$\pi(D \mid \gamma) \propto \lambda_\gamma^{m/2} \exp\left\{-\frac{1}{2}\lambda_\gamma \sum_{p=f,\beta,\mu} \sum_{i\sim j} \left(\gamma_i^{(p)} - \gamma_j^{(p)}\right)^2\right\} = \lambda_\gamma^{m/2} \exp\left\{-\frac{1}{2}\lambda_\gamma \left(\gamma^{(p)}\right)^T W \gamma^{(p)}\right\} \quad (16)$$

where $\sum_{i\sim j}$ is the set of pairwise adjacencies, and matrix W is defined as

$$W_{ij} = \begin{cases} -1 & \text{if } i \text{ and } j \text{ are adjacent} \\ n_i & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

and n_i is the number of neighbors to location i. The prior distribution for the hyperparameter λ_γ in (16) can be chosen as a Gamma distribution

$$\pi(\lambda_\gamma) \propto \lambda_\gamma^{a-1} e^{-b\lambda_\gamma} \quad (18)$$

Finally, the posterior distribution of parameters (γ, λ_γ) given observed data D = (f_o,β_o, μ_o; ψ_o,θ_o, τ_o) can be written as

$$\pi(\gamma, \lambda_\gamma \mid D) \propto L(D \mid \gamma)\pi(\gamma \mid \lambda_\gamma)\pi(\lambda_\gamma). \quad (19)$$

Estimation and inference are based on this posterior distribution. Note that we only need to know the posterior distribution up to a constant proportionality for our Markov chain Monte Carlo simulations discussed in the next section.

I-4. Markov Chain Monte Carlo Simulations

Parameter sampling methods based on Markov chains incorporate the required search aspect in a framework in which it can be proven that the correct distribution is generated, at least in the limit as the length of the chain grows. Writing (γ, λ_γ)$^{(t)}$ for the set of variables at time step t, where γ itself is a vector, the chain is defined by giving an initial distribution for (γ, λ_γ)$^{(0)}$ and the transition probabilities for (γ, λ_γ)$^{(t)}$ given the value for (γ, λ_γ)$^{(t-1)}$. These probabilities are chosen so that the distribution of (γ, λ_γ)$^{(t)}$ converges to that for (γ, λ_γ) as t increases and so that the Markov chain can feasibly be simulated by sampling from the initial distribution and then in succession from the conditional transition distributions.

Typically the Markov chain explores the space in a "local fashion". In some methods for example (γ, λ_γ)$^{(t)}$ differs from (γ, λ_γ)$^{(t-1)}$ in only one component of the state, e.g., it may differ with respect to $\gamma_i^{(t)}$, a component of γ, for some i but have $\gamma_j^{(t)} = \gamma_j^{(t-1)}$ for $j \neq i$. Other methods may change all components at once but usually by only a small amount. Locality is often crucial to the feasibility of these methods. In the Markov chain framework, it is possible to guarantee that such step-by-step local methods eventually produce a sample of points from the globally correct distribution. The procedure implemented in this study can be summarized as follows:

(1) Initialize parameters at some value (γ, λ_γ)$^{(0)}$. Theoretically, these parameters can be initialized using any values taken from the initial distribution. For example, one can initialize (γ, λ_γ) by

drawing a set of random numbers. In this study, we choose γ such that the initial parameter fields computed from γ are close to the mean fields, which were given.

(2) Update each γ_i according to Metropolis rules:

- Draw a value γ_i^* from the uniform distribution $U[\gamma_i^{(t-1)} - r, \gamma_i^{(t-1)} + r]$, where r is a pre-determined small number. Let γ^* be a vector that differs from $\gamma^{(t-1)}$ only in their i^{th} component, i.e., $\gamma^* = (\gamma_1^{(t-1)}, \ldots, \gamma_{i-1}^{(t-1)}, \gamma_i^{(t)}, \gamma_{i+1}^{(t-1)} \ldots, \gamma_n^{(t-1)})^T$.
- Compute $\eta = \pi(\gamma^*, \lambda_\gamma | D) / \pi(\gamma^{(t-1)}, \lambda_\gamma | D)$. Accept the new value γ_i^* with probability $\min(1,\eta)$, or else reject new value γ_i^* (i.e., keep γ_i unchanged). In other words, if the newly proposed value increases the posterior probability (i.e., $\eta > 1$), the new value is accepted. Note that even if the proposed value reduces the posterior probability (i.e., $\eta < 1$), the value could still be accepted with a probability of η.

(3) Update λ_γ given γ according to the following posterior distribution of λ_γ, again using Metropolis rules:

$$\pi(\lambda_\gamma \mid \gamma) \propto \pi(\gamma \mid \lambda_\gamma)\pi(\lambda_\gamma) \sim \Gamma\left(a + \frac{m}{2}, b + \frac{\gamma^T W \gamma}{2}\right), \tag{20}$$

where a and b are two prescribed constants.

(4) Repeat steps 2 and 3 until the chain converges.

To reduce the possible effect of the starting values, results from some early iterations (called burn-in period) are discarded.

I-5. Synthetic Example

In this section, we demonstrate the effectiveness of the approach and examine its performance on a synthetic example in which "data" are generated from a forward run of the numerical model. This synthetic study is being performed in advance of applying the method to the water-content data collected on the Pajarito Plateau. For this hypothetical problem, the flow domain is a soil column with a depth of L = 10 m, uniformly discretized into 100 segments (101 nodes). The pressure head is prescribed at the bottom as $\psi(0) = 0$ (the water table) and water infiltration with a rate of q = 0.002 m/day is prescribed at the top. The statistics of soil properties for this hypothetical soil are given as $\langle f \rangle = 1.0$, $\sigma_f^2 = 1.0$, $\langle \beta \rangle = 0.5$, $\sigma_\beta^2 = 0.01$, $\langle \mu \rangle = -0.9$, $\sigma_\mu^2 = 0.01$, and a correlation length of $\lambda_f = \lambda_\beta = \lambda_\mu = 1.0$. The variability of these parameters can also be given in terms of the coefficient of variation as CV_{Ks} = 131%, $CV_\alpha = CV_n \approx$ 10%. We then generate three random fields as ``true" parameter fields, using the specified statistics and exponential covariance functions for three soil properties. These "true" fields are used as references to access the quality of our inverse model. We solve flow equations (both steady state and transient flow) and transport equations using these true parameter fields to obtain "true" head fields (steady state or transient) and concentration fields (or travel time).

We take $n_f = n_\beta = n_\mu = 5$ samples from these true parameter fields as our direct measurements of f, β, and μ. We also take $n_\psi = n_\theta = 20$ samples for the pressure head and water content and n_C = 6 concentration samples at three elapsed times t = 0.1, 10.0, and 100. After taking all these measurements, we proceed as though the ensemble statistics (the mean, variance, and correlation

lengths) used in generating these original parameter fields are not available, and that all we have are the direct and indirect measurements. Our goal is to estimate three parameter fields using these measurements.

As a first step, we may need to estimate sample statistics of hydrologic properties. Several methods can be used to estimate the sample statistics, i.e., the mean, variance, and correlation length. One simple way is to compute the mean and the variance from direct measurements and find the correlation length by fitting the variogram. An alternative is to estimate these statistics from the maximum likelihood method using both direct and indirect measurements. In our Markov Chain Monte Carlo method (MCMC), these statistics can be estimated simultaneously in the inverse process. However, for simplicity, and also for comparison with some other inverse methods (such as kriging or cokriging methods), in our preliminary study, we compute these statistics from direct measurements only. These estimates are $\langle f \rangle = 1.711$, $\sigma_f^2 = 1.455$, $\langle \beta \rangle = 0.477$, $\sigma_\beta^2 = 0.012$, $\langle \mu \rangle = -0.848$, $\sigma_\mu^2 = 0.008$, and a correlation length of $\lambda_f = \lambda_\beta = \lambda_\mu \approx 1.2$.

For the MCMC method, based on the domain size and the estimated correlation length of about 1.2, we use a grid of 18 basic kernel locations, more-or-less uniformly distributed in the domain, as shown in Figure 1. The error matrices $\sum_p$ are chosen to be $\varepsilon_p^2 I_{n_p}$, where $p = f, \beta, \mu, \psi, \theta$, or C, and I_n stands for an identical matrix of $n \times n$, and ε's are prescribed standard deviations for errors of variables p's. Here we choose $\varepsilon_f = 0.10$, $\varepsilon_\beta = 0.02$, $\varepsilon_\mu = 0.02$, $\varepsilon_\psi = 0.01$, $\varepsilon_\theta = 0.005$, and $\varepsilon_C = 0.002$.

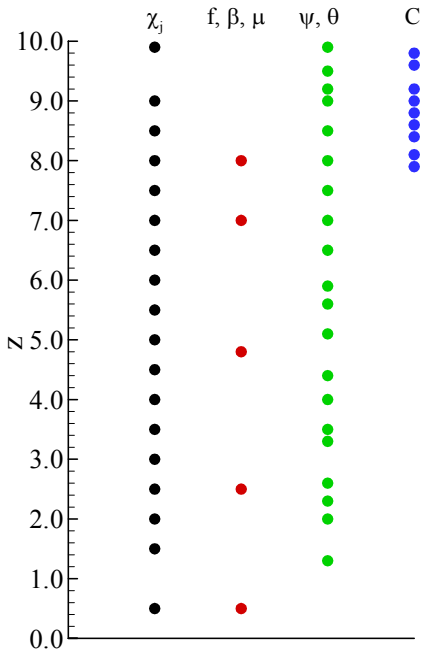

Figure 1. Layout of the problem configuration for the hypothetical example.

The estimated hydrologic parameter fields from the MCMC method (blue curves) are illustrated in Figure 2, as compared to the true parameters fields (red curves). The results demonstrate that the estimated parameter fields match the true ones very well. Note that the estimated values at the conditioning points deviate from their corresponding true values, because the specified measurement errors in the MCMC method allow the estimated values vary within the range of the measurement error. The degree of such deviations is characterized by the standard deviation of errors specified by ε_p, where $p = f$, β, or μ.

Figures 3, 4, and 5 compare the true pressure head, moisture content and concentration profiles against the simulated ones at three different times. The results suggest that the MCMC method is capable of providing excellent fits to the hydrologic data, making it an appropriate method for the water-content data for the Pajarito Plateau.

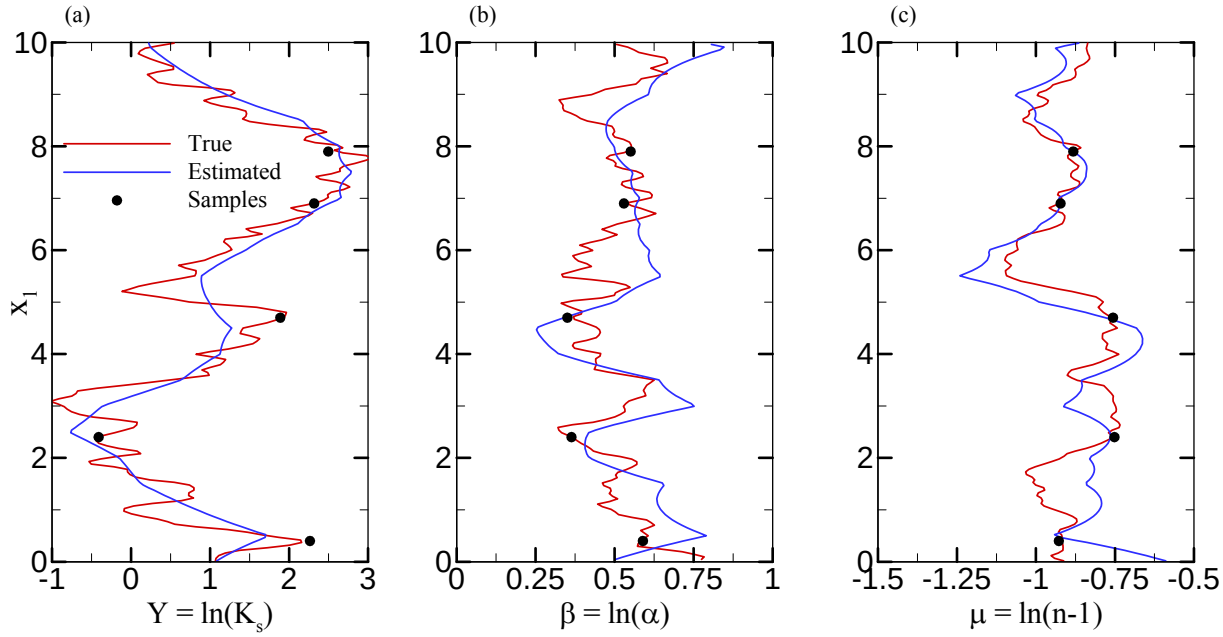

Figure 2. Comparison of the true soil properties and inverse results for the synthetic example.

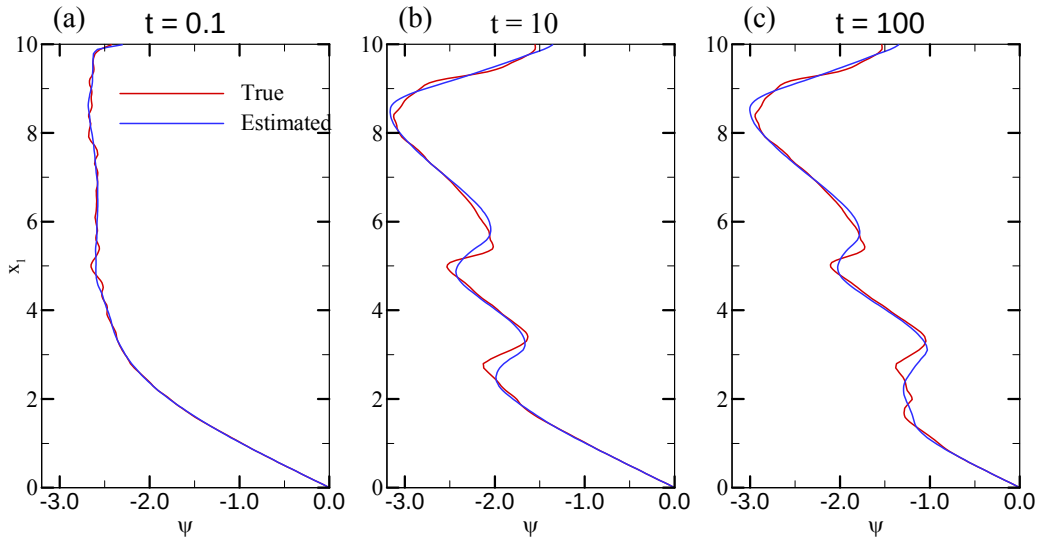

Figure 3. Comparison of the true and simulated pressure head profiles for the synthetic example.

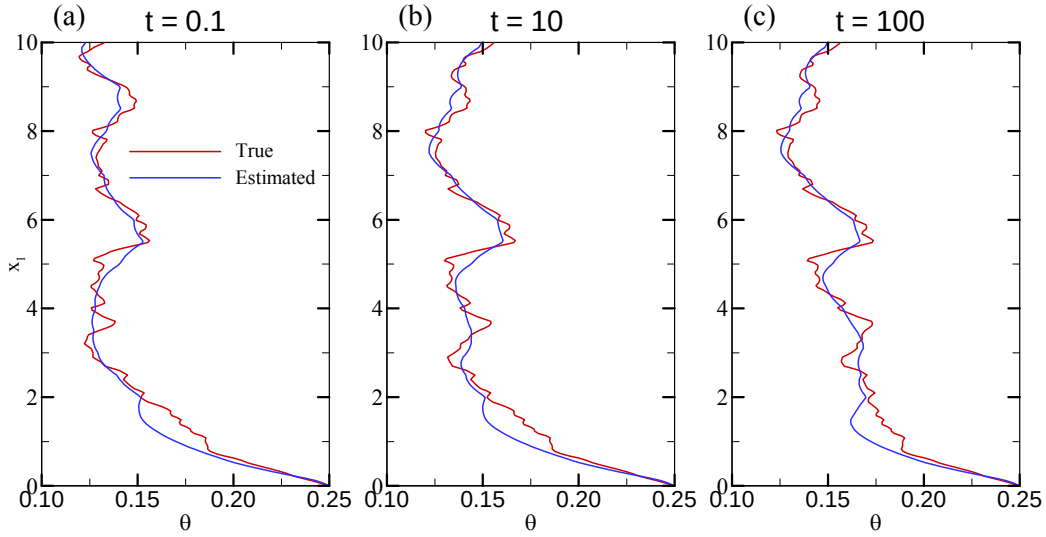

Figure 4. Comparison of the true and simulated moisture content profiles for the synthetic example.

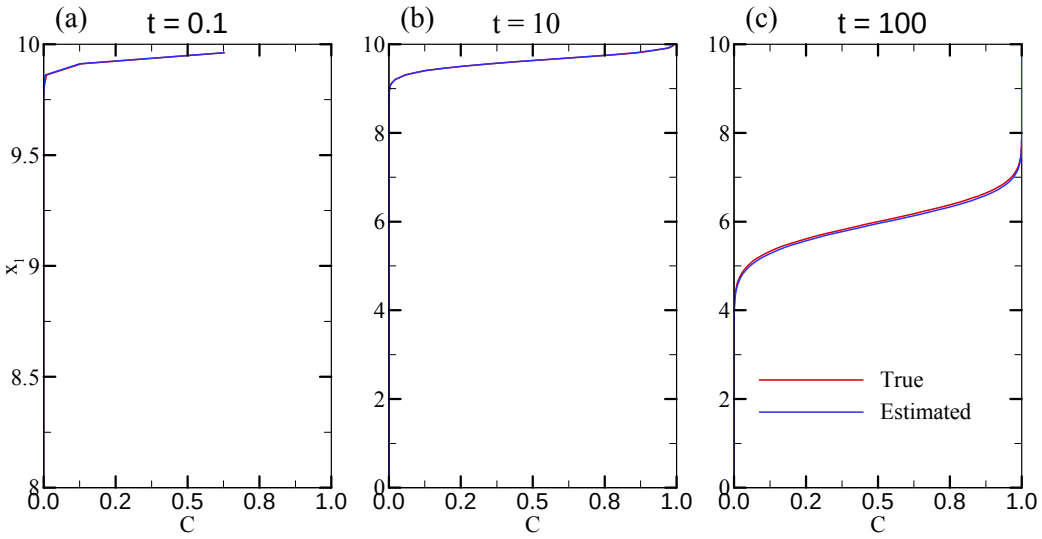

Figure 5. Comparison of the true and simulated concentration profiles for the synthetic example.

II. Application of the MCMC Method to LANL Water Content Data

In this section, we apply the MCMC method described above to use the water-content data to estimate local infiltration rate in the vicinity of a number of wells at the LANL site. Since there is significant model development required to implement this new analysis approach, we begin by discussing the model setup. Then, we examine one of the wells, MCOBT-4.4, in greater detail, using it to assess the effectiveness of the technique. We then present the analyses for the other wells in a more concise way.

Table 1 lists the hydrologic properties used for the model. The table contains both the permeability and porosity values used for each unit, as well as unsaturated hydraulic parameters, α and n, for defining the van Genuchten (1980) constitutive relationship. We assume that porosity is a deterministic constant (a constant for each type of stratigraphic unit), while the permeability, pre-size distribution parameter α, and fitting parameter n, are spatially random functions and are modeled by log-normal distributions. The property values of a unit listed in the table will be used as initial values for the MCMC simulation if the unit occurs in a borehole.

Table 1. Stratigraphic Units and Their Hydraulic Properties at the Los Alamos Area

No	Description	k(m^2)	α (1/m)	n	porosity
1	Tsfu (all units below Tsfuv, undifferentiated, volcaniclastic)	2.65E-13	5	2.68	0.35
2	Tb1A (Cerros del Rio basalt-occurs within Santa Fe group)	2.96E-13	5	1.5	0
3	Tsfuv (Santa Fe Group, aquifer unit)	2.65E-13	5	2.68	0.35
4	Tb2 (Cerros del Rio basalt-occurs within Tsfuv)	2.96E-13	5	1.5	0
5	Tpt (Puye Formation, Totavi equivalent)	4.73E-12	5	2.68	0.35
6	Tpf (Puye Formation, fanglomerate)	4.73E-12	5	2.68	0.35
7	Tb4 (Cerros del Rio basalt-occurs within Puye Formation)	2.96E-13	5	1.5	0
8	Tt1 (Tschicoma dacite-occurs within Puye Formation)	2.96E-13	5	1.5	0
9	Tt2 (Tschicoma dacite-occurs within Puye Formation)	2.96E-13	5	1.5	0
10	Qbog (The Guaje pumice bed)	1.53E-13	0.081	4.026	0.667
11	Qbof (Otowi member of Bandelier Tuff)	7.25E-13	0.66	1.711	0.469
12	Qct (Cerro Toledo interval)	8.82E-13	1.52	1.506	0.473
13	Qbtt (Basal Pumice Unit, Tshirege member of Bandelier Tuff)	1.01E-12	1.52	1.506	0.473
14	Qbt1g(Glassy unit, Tshirege member of Bandelier Tuff)	3.68E-13	2.22	1.592	0.509
15	Qbt1v(Vitric unit, Tshirege member of Bandelier Tuff)	1.96E-13	0.44	1.66	0.528
16	Qbt2 (Unit 2, Tshirege member of Bandelier Tuff)	7.48E-13	0.66	2.09	0.479
17	Qbt3 (Unit 3, Tshirege member of Bandelier Tuff)	1.01E-13	0.29	1.884	0.469
18	Qbt3t (Unit 3t, Tshirege member of Bandelier Tuff)	5.10E-13	2.57	1.332	0.466

19	Qbt4 (Unit 4, Tshirege member of Bandelier Tuff)	9.18E-14	0.667	1.685	0.478
20	Qbt5 (Unit 5, Tshirege member of Bandelier Tuff)	1.43E-14	0.17	1.602	0.349

II-1. Model Setup

Although the method is not restricted to these assumptions, the simplified model used to perform the MCMC analyses of the moisture content profiles is a one-dimensional, steady state flow model with uniform numerical grids. The goal of the analysis is to estimate the infiltration rate and hydrologic parameters associated with the measured water contents from the wells. Therefore, in general, the infiltration rate, the saturated hydraulic conductivity, and the van Genuchten parameters are jointly varied in the inverse model runs. For testing purposes, other strategies are employed to examine the influence of these choices on the inversion results.

FEHM is the numerical model used to perform the model runs. Grid spacing was chosen so that, in general, the spacing was smaller than the spacing of water content measurements. This approach reduces the possible loss of conditional points. For cases in which more than one measurement is located in single computational grid cell, the average water content value of all measurement points were taken this node as a new conditional point. In addition, since for each node there are three parameters to be estimated, there are practical limitations to the resolution of the grid for situations in which parameters are allowed to vary on a node-by-node basis. For these reasons, we selected a grid spacing ranging from about 3 to 5 feet, depending on the problem.

The one-dimensional column representing a given well is divided into a number of zones, mainly based on the hydrostratigraphy determined from well logs. When detailed information (upper and lower bounds) about subdivisions in a particular formation is available, each subdivision is defined as an individual zone. Initially, hydraulic properties are assumed to be the same for all subdivisions of the formation, and the MCMC method seeks to fit the water content data by adjusting the hydrologic parameters in the column. Three strategies are employed for this purpose:

1. Properties are assumed to be uniform random constants within a zone, but vary from zone to zone.
2. Properties are defined by a set of kernel functions, the coefficients of which are updated sequentially.
3. Properties are defined on a node-by-node basis.

These methods represent increasingly complex models of the heterogeneities in hydrologic properties. Part of our study will consist of assessing which method is the appropriate level of complexity for a given data set, taking into consideration data sparseness and computational efficiency. In most cases, we will not use the node-by-node approach (method 3) due to computational costs and data limitations. For example, for a soil column of 100 nodes, there are 300 hydraulic parameters plus the infiltration rate, which means that for any particular parameter, 301 model runs are required to perform an update. The reason to develop method two

is that it seems like a reasonable compromise that allows heterogeneities within thick units to be modeled without an excessive number of parameters.

Finally, because the water content measurements are usually available in the upper part of the well and hydraulic properties in the deep zones that are far away from the measurement locations have little impact on data fitting, properties in these deep zones are modeled as random constants.

In each MCMC simulation, the run starts from an initial set of soil parameters and infiltration rate, and the parameter values are updated sequentially based on the rules described in Section I. One of the virtues of the approach is that the parameter uncertainty statistics can be derived from the variability of values obtained during the chain. However, the estimated parameters in the first part of the chain strongly depend on the choice of initial settings. Thus, this initialization phase of the simulation (called the burn-in period) is ignored when computing parameter statistics. The length of burn-in, which is problem dependent, is determined graphically from a plot of the negative log posterior versus the number of updates. Initially, this metric is possibly very large, since the initial setting may significantly deviate from the true solution. The burn-in period is approximated graphically as the point in the simulation at which the negative log posterior decreases to the point at which the variation becomes small.

Since the MCMC method takes a perturbation approach to parameter updating (i.e., the new value for any parameter is derived by adding a possible perturbation to the old value), the sequence of values for each parameter are not completely independent. As a consequence, the computed parameter variance may be artificially small. Because we seek to use the analysis to compute the statistics (mean and variance) for each parameter, we take a subset of the sequence by selecting values within a predefined interval. The variance for parameter is then used to construct the confidence intervals around the mean predictions.

In the following subsections, we will illustrate our inversion results for the selected wells on the LANL site. First, we examine well MCOBT-4.4 in detail to investigate the effect of different algorithmic choices on the inversion results.

II-2. Well MCOBT-4.4

MCOBT-4.4 is an intermediate well in Mortandad canyon. Geologic units encountered in MCOBT-4.4 are, in descending order: canyon-bottom alluvium; deposits of the Cerro Toledo interval; the Otowi Member of the Bandelier Tuff, including the basal Guaje Pumice Bed; an upper sequence of fanglomerate and sand deposits of the Puye Formation; lavas, interflow units, and subflow deposits of the Cerros del Rio volcanic field; and a lower sequence of fanglomerate deposits of the Puye Formation. Depths and elevations of the contacts between these units are shown in Figure 6. Canyon-bottom alluvium (Qal) was cored from 0 to 63.7 ft depth at MCOBT-4.4. The alluvium consists predominantly of moderately weathered detritus of the Tshirege Member of the Bandelier Tuff, and is unconsolidated. In this study, we exclude this layer of alluvium in our simulations. The simulation domain ranges from the depth of 64ft (elevation 6769.2ft, or 2063.25m) to 493ft (elevation 6343.2ft, or 1933.4m), where perched water occurs. The column is uniformly discretized into 143 elements of size 3ft (0.914m). There are 30

measurements of water content available. Here we assume that the error matrix for the water content as appeared in (15) is a diagonal $\sum_{\theta} = \varepsilon_{\theta}^2 I_{n_\theta}$ with $\varepsilon_{\theta} = 0.01$.

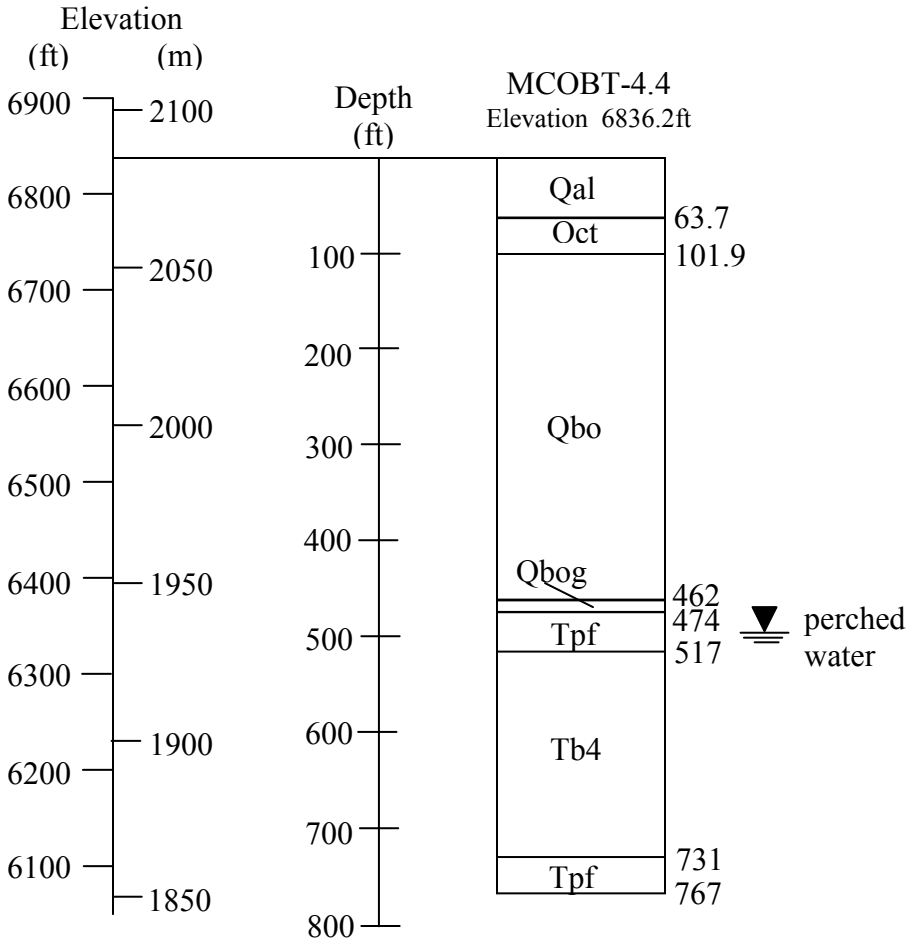

Figure 6. Stratigraphy at MCOBT-4.4

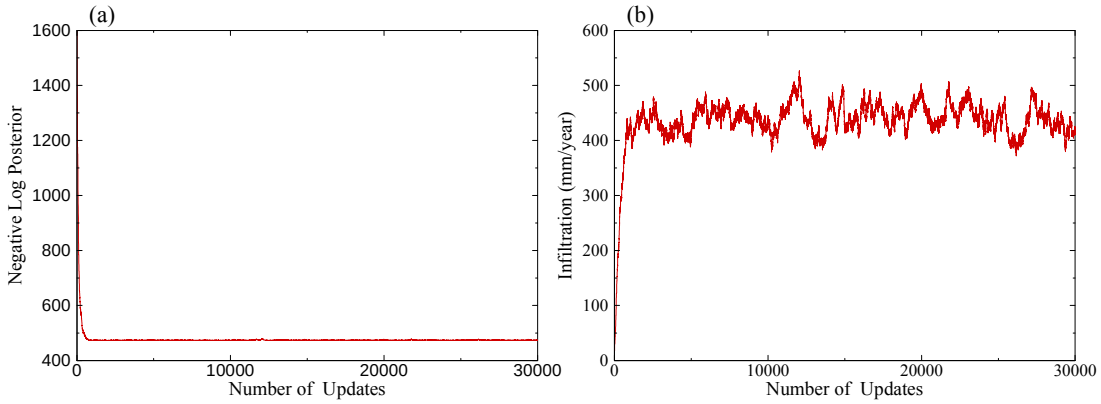

Figure 7. The negative log posterior (a) and the infiltration rate (b) derived using deterministic hydrologic properties and random infiltration rate (starting from a low value)

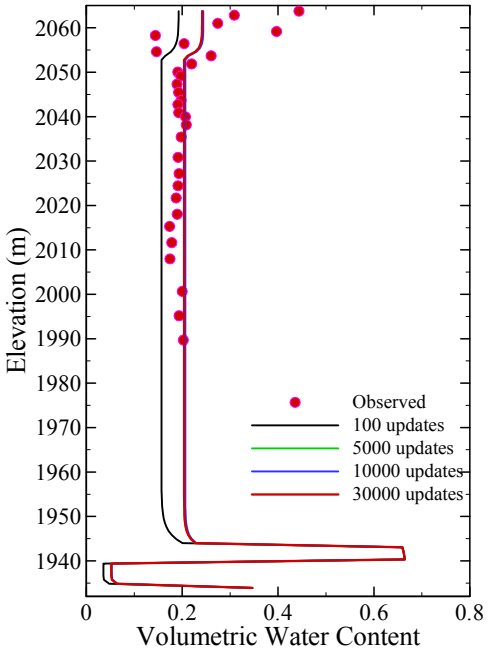

Figure 8. Comparison of observed water content with fitted profiles of different numbers of updates for the case with deterministic hydrologic properties and random infiltration rate (starting from a low value)

In the first case, we fix the hydrologic properties of the column based on stratigraphic units at the well, while allowing the infiltration rate to vary, starting from a lower value of 5×10^{-7} kg/sec (=15.7 mm/year) with a possible maximum increment of r = 2×10^{-7} kg/sec at each update. The infiltration rate and its corresponding negative log posterior function, which measures the closeness of modeled and observed water content, are shown in Figure 7. Figure 7(a) shows that after a few hundred updates the model becomes stabilized with a negative log posterior of 474, which is equivalent to a root-mean-square-error (RMSE) of water content = 0.056. The error is

are relatively large fraction of the water content. As illustrated in Figure 8, the fit between the observed and modeled water content is not very good for this run. With only the infiltration rate being varied, and no adjustment of the hydrologic properties, the method is limit in its ability to capture the details of the water content profile. Regarding the infiltration rate estimation, after ignoring the first 1000 realizations of the chain, the statistics of the infiltration rate are a mean infiltration rate of $\langle q \rangle = 442.6$ mm/year and a standard deviation of $\sigma_q = 25.6$ mm/year.

Next, we examine the results with an identical parameter strategy as above, but start the MCMC simulation with a very high infiltration rate of 5×10^{-5} kg/sec (=1568 mm/year) and a possible maximum increment of $r = 1\times10^{-6}$ kg/sec at each update. The resulting Markov chain is illustrated in Figure 9. The mean and standard deviation of the infiltration rate after ignoring the first 1000 realizations are: $\langle q \rangle = 439$ mm/year and $\sigma_q = 22.4$ mm/year. Thus, the method has the desirable property that final infiltration rate is independent of the initial value. However, it seems from these two cases that good fits to the measured water content data cannot be obtained by simply varying the infiltration rate alone. Modification of hydrologic parameters appear to be required to fit the observed water-content data.

When conducted a modeling exercise such as this, it is important to note that the inverse problem is, in general, ill-posed, and the solution to the problem is non-unique. Therefore, the ability to impose constraints on parameter variability based on *a priori* knowledge or expert judgment is critical for obtaining realistic infiltration estimates. The MCMC method allows this to be done very easily. In the calculations that follow, when updating hydrologic properties k, α, and n, we set the following constraints for each stratigraphic unit: $10^{-2}\, k_0 \le k \le 10^2\, k_0$, $0.5\,\alpha_0 \le \alpha \le 2\alpha_0$, and $\max(0.5n_0, 1.05) \le n \le 2n_0$, where k_0, α_0, and n_0 are respectively permeability, α, and n values for the stratigraphic unit.

In the third case, in addition to the variable infiltration rate, we allow the hydraulic properties to vary as random constants for the three parameters of each unit, but the medium is assumed to be uniform within the layer (method 1 in Section II-1). The results are illustrated in Figures 11-13. Note that although the negative log posterior becomes stabilized after a few hundred updates, the infiltration rate can vary significantly (Figure 11b). Any of these infiltration rates, together with a combination of variable hydrologic properties, can fit the water content data equally well. Finally, note that the fits, though improved, still exhibit systematic deviations and bias (Figure 12).

These results suggest that some of the layers at well MCOBT-4.4 need to be modeled by sets of kernel functions (method 2) or on a node-by-node basis (method 3) to provide the level of heterogeneity in hydrologic properties required to obtain a better match to the profile.

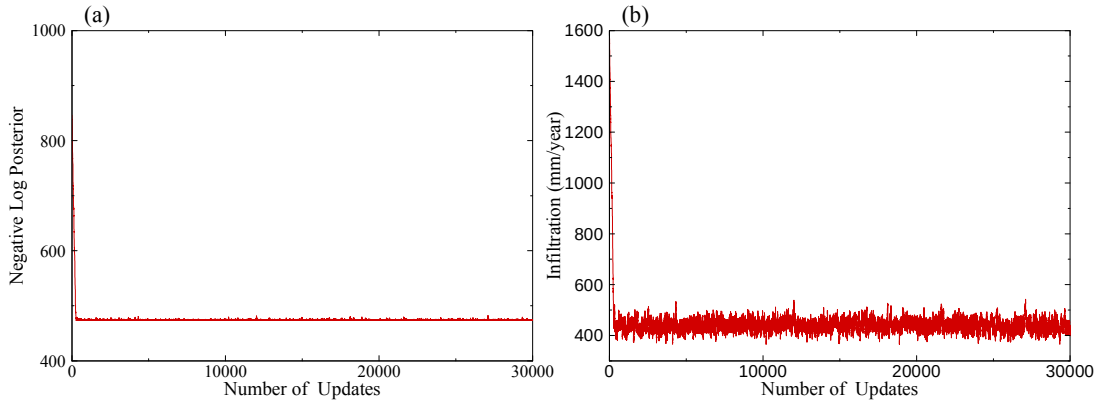

Figure 9. The negative log posterior (a) and the infiltration rate (b) derived using deterministic soil properties and a random infiltration rate (starting from a high value)

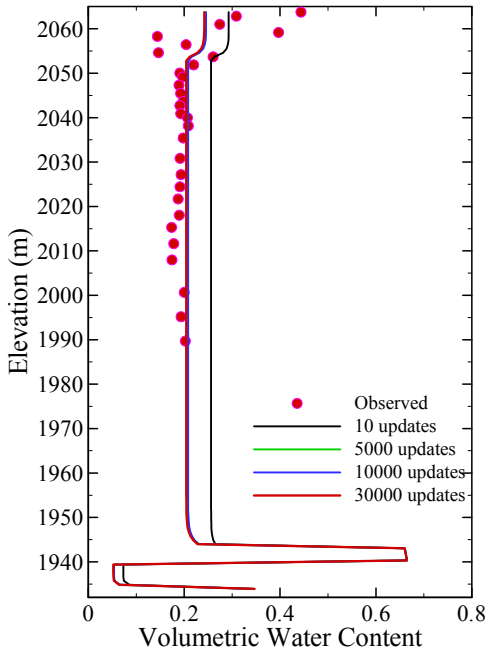

Figure 10. Comparison of observed water content with fitted profiles of different numbers of updates for the case with deterministic soil properties and a random infiltration rate (starting from a high value)

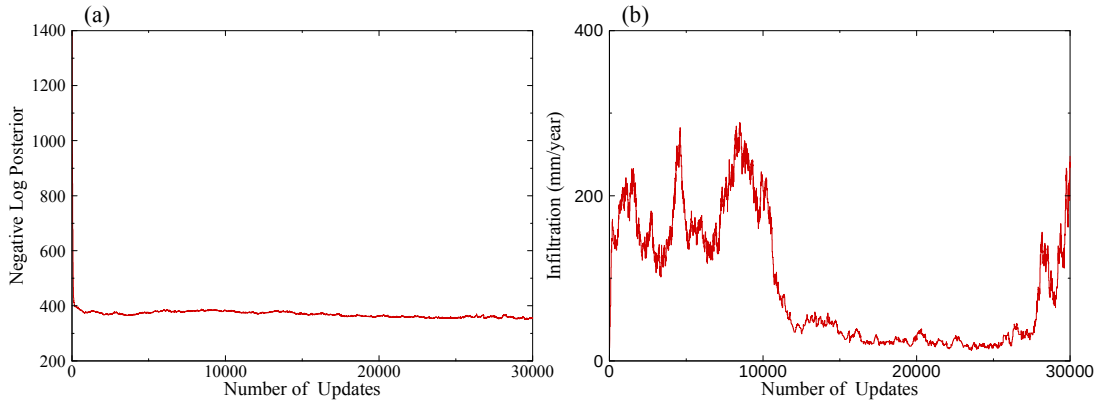

Figure 11. The negative log posterior (a) and the infiltration rate (b) derived using random constant soil properties and a random infiltration rate.

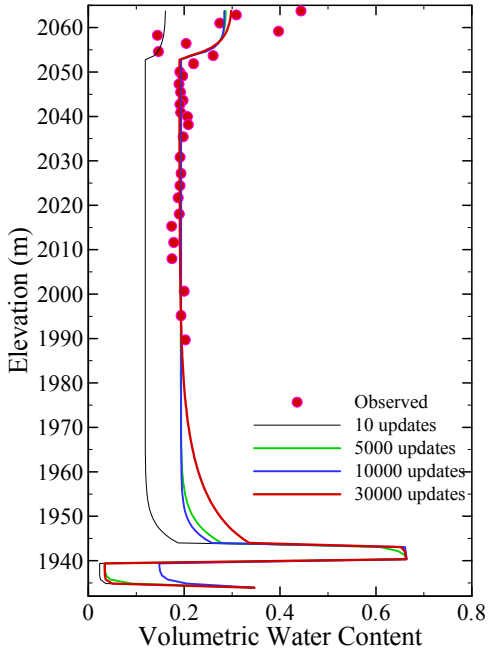

Figure 12. Comparison of observed water content with fitted profiles of different number of updates for the case with random constant soil properties and a random infiltration rate.

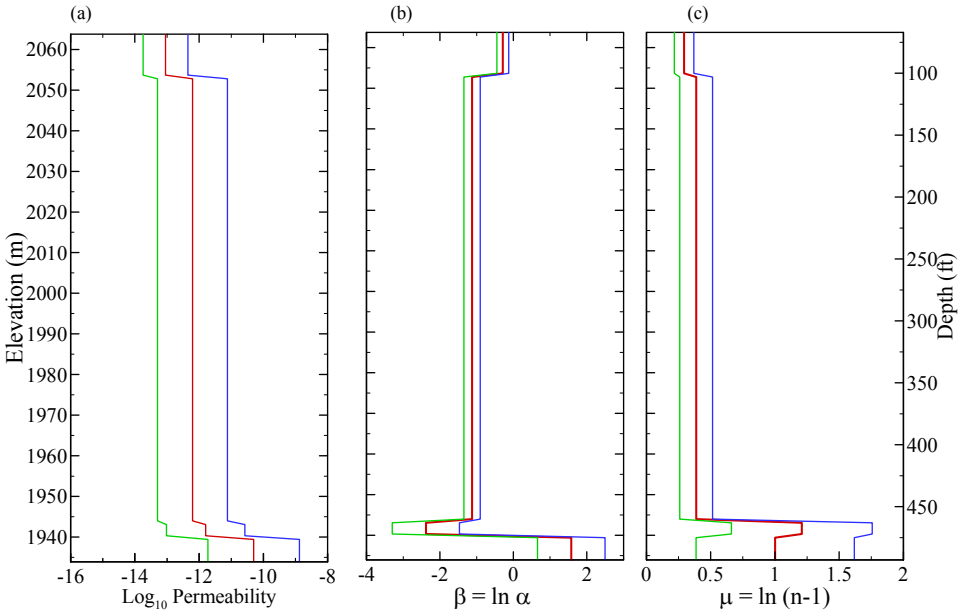

Figure 13. Estimated statistics of hydrologic properties, mean (red curves) and plus/minus two standard deviations (blue and green curves), for the case with random constant soil properties and a random infiltration rate.

In the next run (case 4), we model the soil properties in Qbo (depth 101.9 to 462 ft) by 10 kernel functions while other three thin layers (Qct, 64-101.9ft, Qbog, 462-474ft, and Tpf, 474-493 ft) are modeled by individual nodes. This strategy results in a total of 143 grid nodes, 23 of which are modeled individually, and 121 nodes of which are represented by 10 kernel functions. This yields 33 groups of hydrologic parameters, plus one additional parameter for the infiltration rate. The initial hydrologic properties at grid nodes are assigned based on their stratigraphic units, and the initial infiltration rate is 5×10^{-7} kg/sec (=15.7 mm/year).

Figure 14(a) depicts the negative log posterior as a function of the number of updates. The figure shows that the negative log posterior drops quickly at the beginning as the simulation proceeds, and then stabilizes at about 10, which is equivalent to the root-mean-square-error (RMSE) of 0.0082 for water content. Figure 14(b) illustrates the decrease of the negative log posterior for the first few sets of updates as a function of 34 parameter groups, where parameter zero corresponds to the infiltration rate. The figure clearly shows that updates on the infiltration rate have a significant effect on reducing the negative log posterior. On the other hand, variations of the hydrologic properties in the bottom part of the column have a minimal impact on the data fits.

The comparison between the modeled and observed water content is shown in Figure 15(a), for different numbers of updates. The figure indicates that, by updating hydrologic properties at each node the upper part of column, we are able to fit the observed water content values very well. Furthermore, note that in the lower part of Qbo the water content can vary significantly, a consequence of the fact that there is no data in this part of the column. Finally, the water content

in the bottom two layers (below 1945m) changes very little, even though the soil properties in these layers are highly heterogeneous (Figure 16).

The estimated mean and standard deviation of the infiltration rate are 125.3 mm/yr and 30.9 mm/yr, respectively. Because the infiltration rate is normally distributed (Figure 17d), these statistics mean that at the 95 confidence level the actual infiltration rate ranges from 63.5 mm/yr to 187.1 mm/yr.

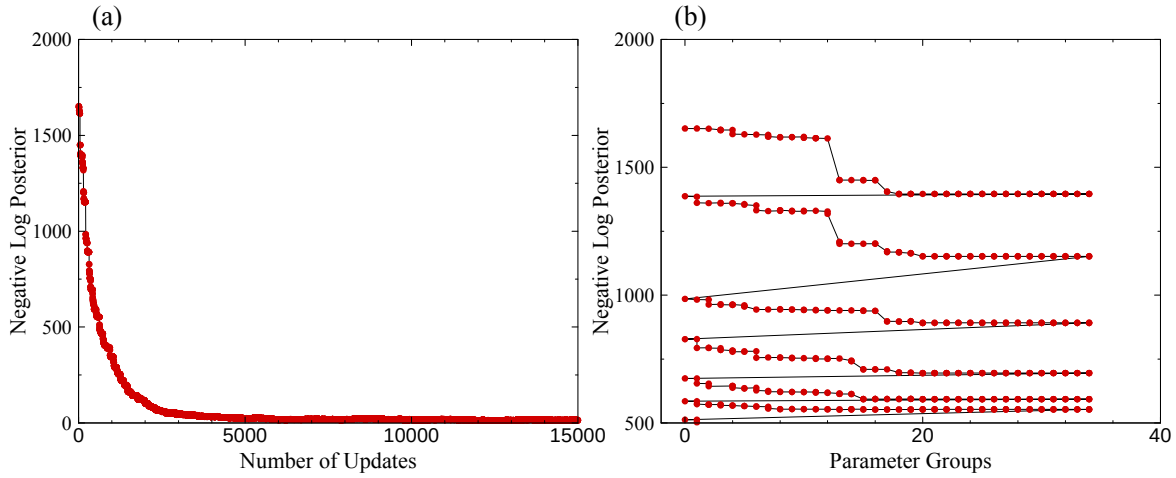

Figure 14. Negative log posterior vs (a) the number of updates, and (b) parameter groups.

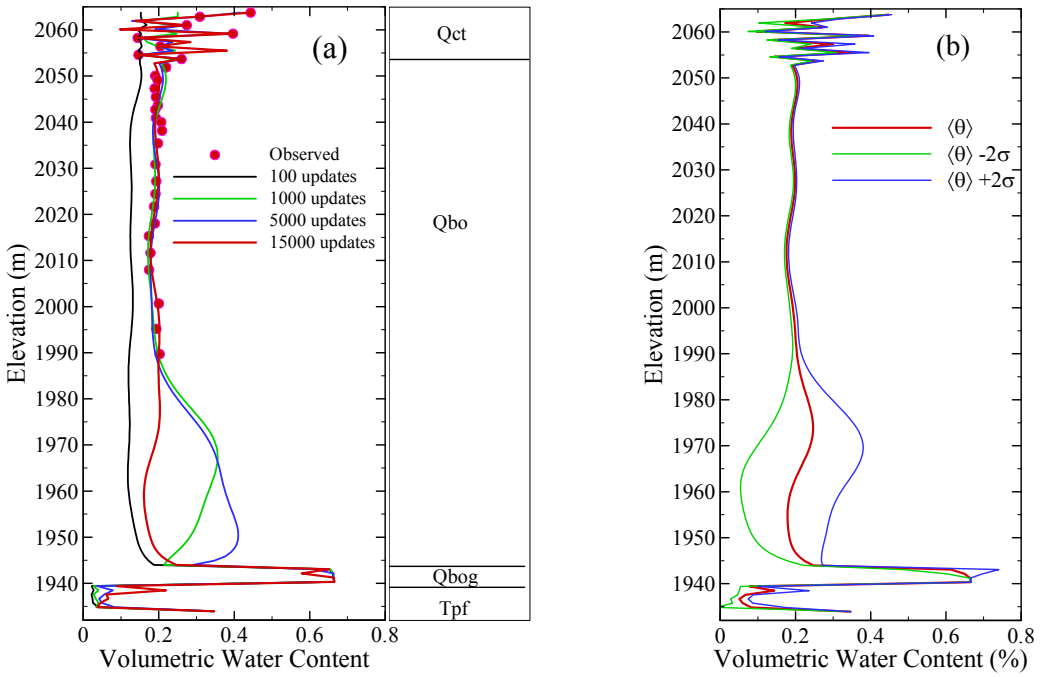

Figure 15. (a) Comparison of observed water content with fitted profiles of different numbers of updates for the case 4, and (b) the mean and confidence interval for volumetric water content.

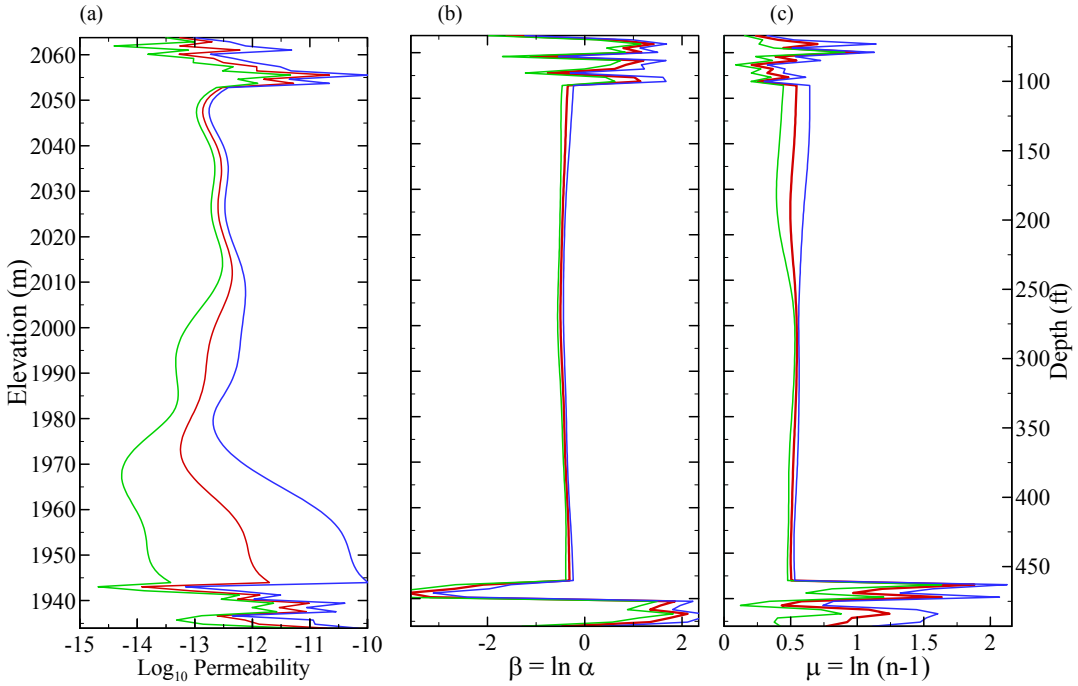

Figure 16. Estimated statistics of hydrologic properties, mean (red curves) and plus/minus two standard deviations (blue and green curves), for the case with random constant soil properties and a random infiltration rate.

All simulations presented thus far have used the mean values of parameters listed in Table 1 for the initial values of the MCMC run. To test the sensitivity of the inversion results on initial hydrologic properties, three more cases were performed with different starting parameter values than those in case 4, which is considered to be the base case. In case 5, we initialize the model with permeability values 20 times larger than the values based on the stratigraphic units. In cases 6 and 7, the initial α and n fields are respectively 1.5 larger than those in case 4.

The estimated mean and standard deviation of infiltration rate for these different cases (cases 4-7) are listed in Table 2, along with the root-mean-square-error (RMSE) for each case. Although the statistics of the infiltration rate differ from case to case, all four cases (the base case and the three sensitivity runs) fit the observed water content equally well, as evidenced from the RMSE in Table 2. The histograms of the infiltration rate for these cases are illustrated in Figure 17. For comparison purposes, the histogram for the base case is also shown. It is interesting to see that the histogram of the infiltration rate for the base case and the case with an initially high permeability setting follows approximately the normal distribution, whereas the cases with high α and n values the histogram exhibit skewness and, in the case of Figure 17b, bimodality.

Table 2. Test Cases for Well MCOBT-4.4

Case #	Description	Mean infiltration rate (mm/yr)	Std deviation of infiltration rate(mm/yr)	RMSE of water content	Comments
1	Deterministic hydrologic properties; random infiltration rate (starting from lower q)	442.6	25.6	0.0562	Poor fitting
2	Deterministic hydrologic properties; random infiltration rate (starting from higher q)	439.4	22.4	0.0562	Poor fitting
3	Random constant hydrologic properties; random infiltration rate	122.9	78.7	0.0488	Normally distributed infiltration rate
4	Correlated hydrologic properties; random infiltration rate, starting from mean properties (base case)	125.3	30.9	0.0084	Normally distributed infiltration rate
5	Correlated hydrologic properties; random infiltration rate, starting from higher α values (1.5 times)	44.3	11.8	0.0089	
6	Correlated hydrologic properties; random infiltration rate, starting from higher n values (1.5 times)	78.4	21.3	0.0098	
7	Correlated hydrologic properties; random infiltration rate; starting from higher Ks values (20 times)	129.2	28.2	0.0090	Normally distributed infiltration rate

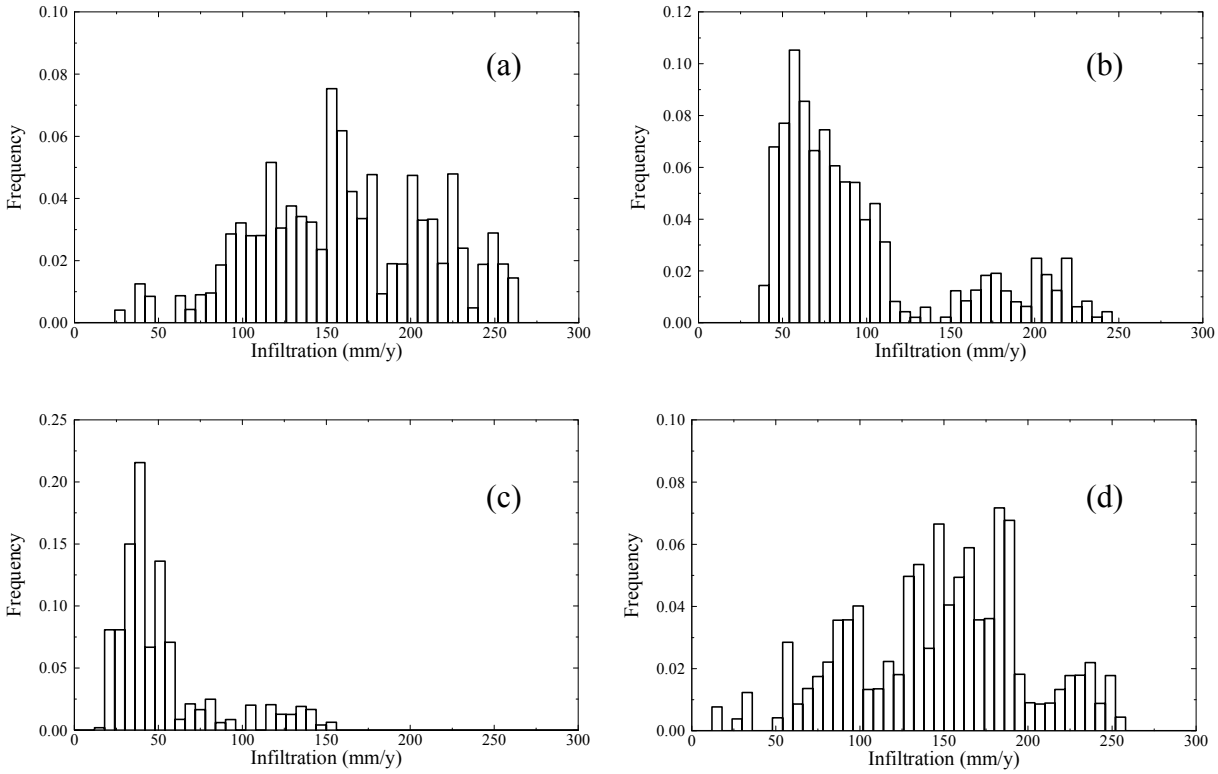

Figure 17. Histograms of the infiltration rate derived from simulations with various initial settings that are different from those based on rock units: (a) permeability is 20 times larger, case 5; (b) α values are 1.5 times larger, case 6; (c) n values are 1.5 times larger, case 7; and (d) based on rock units, case 4, the base case.

II-3. Borehole MCOBT-8.5

For this and the remaining wells analyzed below, the same method used in the base case model was used to fit the measured water content data. Specifically, the infiltration rate and the hydrologic parameters are simultaneously varied in the MCMC inversion process. This allows the results of the different wells to be directly compared.

MCOBT-8.5 is an intermediate well in Mortandad canyon. Geologic units encountered in MCOBT-8.5 consist of the following, in descending order: canyon-bottom alluvium; an interval of canyon-bottom alluvium plus colluvium; the Otowi Member of the Bandelier Tuff including the basal Guaje Pumice Bed; an upper sequence of Puye Formation fanglomerates; lavas, interflow units, and subflow deposits of the Cerros del Rio volcanic field; and a lower sequence of fanglomerate deposits of the Puye Formation (Figure 18). There are 14 water content measurements used in the inverse process. The estimated mean and standard deviation of the infiltration rate (see Table 3) are 21.9 mm/yr and 5.7 mm/yr, respectively, which are much lower

than those at Well MCOBT-4.4. This result is consistent with a relatively dry condition at this location, which agrees with the fact that there is no evidence of perched water at this borehole.

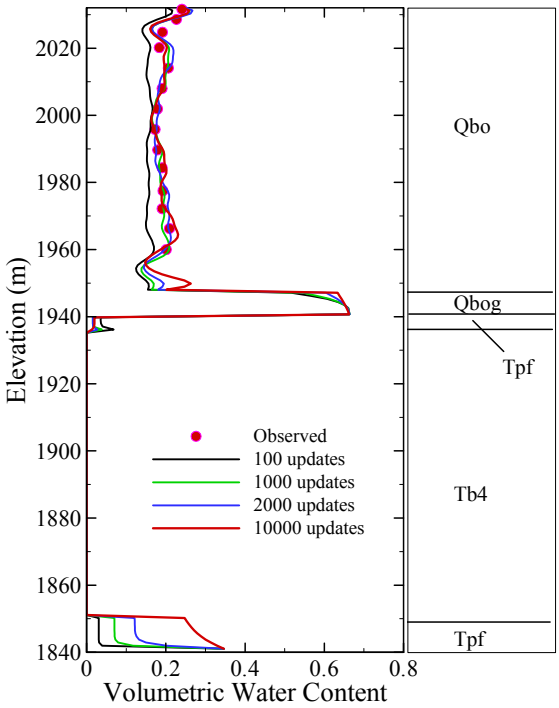

Figure 18. Comparison of observed water content with fitted profiles of different numbers of updates for well MCOBT-8.5.

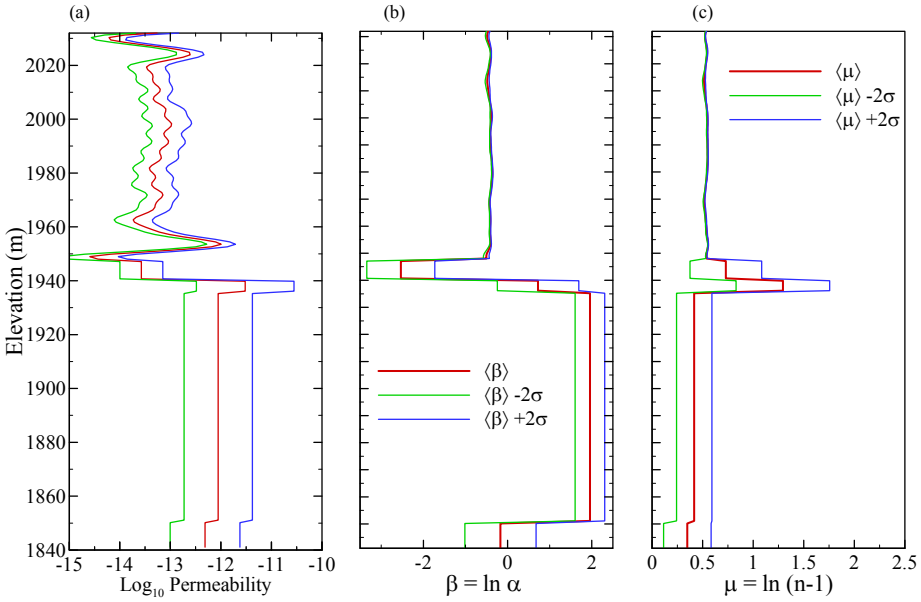

Figure 19. Estimated statistics of hydrologic properties, mean (red curves) and plus/minus two standard deviations (blue and green curves) for well MCOBT-8.5.

II-4. Borehole LADP-3

Borehole LADP-3 is located in Los Alamos Canyon. LADP-3 penetrated a thick sequence of slope-derived colluvium and stream-derived alluvium on the canyon floor before entering bedrock [Broxton et al., 1995]. Bedrock units penetrated by the borehole include the Otowi member of the Bandelier Tuff and gravels of the Puye Formation (Figure 20). The borehole encounters two perched water zones, one in the canyon's alluvium and the other at a depth of 325 ft in the Guaje Pumice Bed. The hydrologic properties and the infiltration rate are estimated using 45 water content measurements. The fitting of modeled water content to the observed is illustrated in Figure 20, and the predicted soil properties and their associated uncertainties are depicted in Figure 21. The estimated mean and standard deviation of the infiltration rate are 220 mm/yr and 16.7 mm/yr, respectively.

II-4. Borehole LADP-4

Borehole LADP-4 is located in DP Canyon. Because of the topographic setting, this region is expected to have lower infiltration rates than the adjacent well LADP-3. LADP-4 penetrated alluvium on the canyon floor and then entered bedrock unit: the Tshirege Member of the Bandelier Tuff (including the Tsankawi Pumice Bed), fluvial sediments of Cerro Toledo interval, the Otowi Member of the Bandelier Tuff (including the Guaje Pumice Bed), and Puye Formation (Figure 22). One purpose of drilling this borehole was to determine if perched groundwater occurs beneath DP Canyon. No perched groundwater was encountered. Moisture contents were determined for 139 samples, but 125 measurements are used in our inverse modeling because the rest of measurements were located in the alluvium, which is excluded from the inverse modeling. The fitting of modeled water content to the observed is shown in Figure 22, and the predicted soil properties and their associated uncertainties are depicted in Figure 23. The estimated mean and standard deviation of the infiltration rate are 10 mm/yr and 1.9 mm/yr, respectively. As expected, the infiltration estimates are more than an order of magnitude lower than those at LADP-3.

It is interesting to notice that, based on the presence of the Guaje Pumice Bed (Qbog), the water content profile should have a peak in this pumice, as shown from the earlier part Markov chain.

II-5. Well R-9

R-9 is located in Los Alamos Canyon near the eastern boundary of Los Alamos National Laboratory. The principal geologic units encountered in R-9, in descending order, consist of alluvium, basalts of the Cerros del Rio volcanic field and basaltic tephra, old alluvium, sediments of the Puye Formation (see Figure 24). In the model, we have excluded the alluvium and modeled only two large layers, the basalt and Puye Formation. Because of large variations on measured water contents, we further divide the basalt into five subunits based on the petrographic and chemical character of the unit [Broxton et al., 2001]: Upper Tholeiite, Lower Tholeiite, Upper Alkalic Basalt, Lower Alkalic Basalt, and Basaltic Tephra, and the Puye Formation into two subunits: Upper Puye Formation and Lower Puye Formation. The inversion

results are shown in Figures 24 and 25. The estimated mean and standard deviation of the infiltration rate are 139.9 mm/yr and 84.2 mm/yr, respectively.

II-6. Well LAOI(A)-1.1

The stratigraphic units at this borehole and the inversion results are illustrated in Figures 26 and 27. The estimated mean and standard deviation of the infiltration rate are 523.7 mm/yr and 45.9 mm/yr, respectively. This infiltration rate is the highest of any well examined

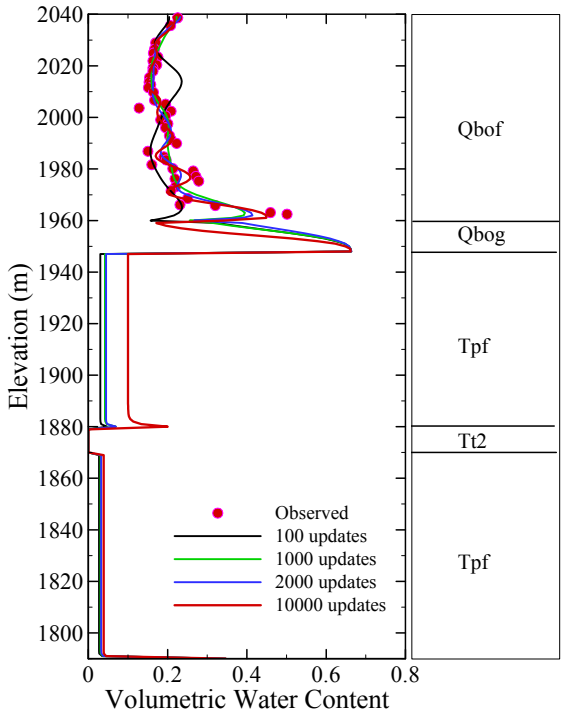

Figure 20. Comparison of observed water content with fitted profiles of different numbers of updates for well LADP-3.

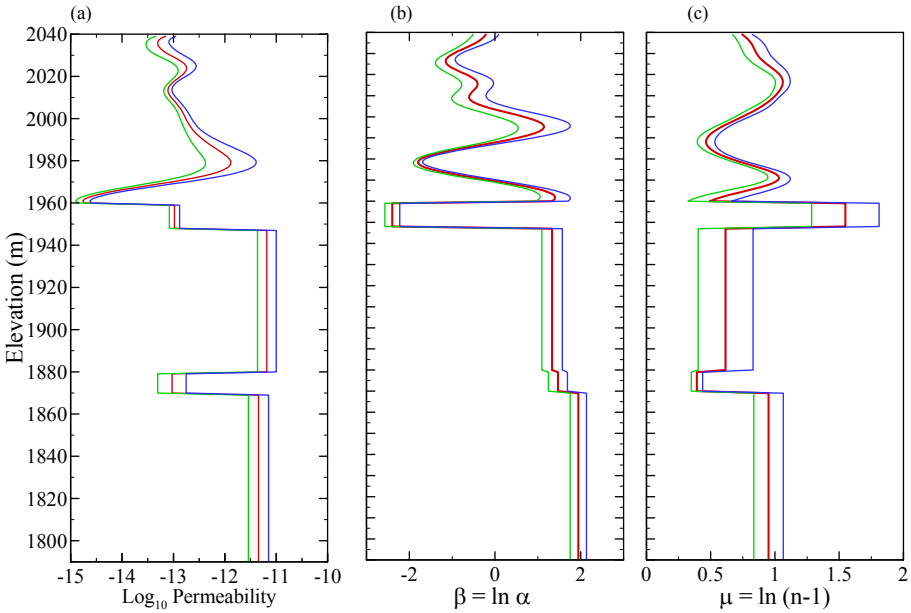

Figure 21. Estimated statistics of hydrologic properties: mean (red curves) and plus/minus two standard deviations (blue and green curves) for well LADP-3.

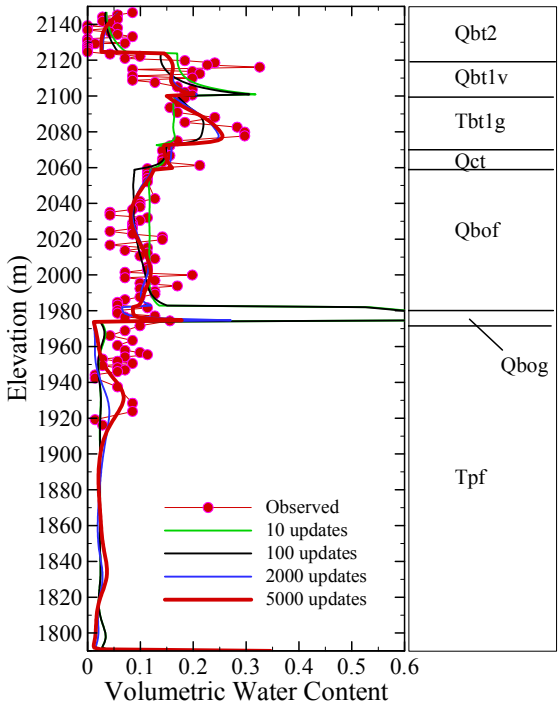

Figure 22. Comparison of observed water content with fitted profiles of different numbers of updates for well LADP-4.

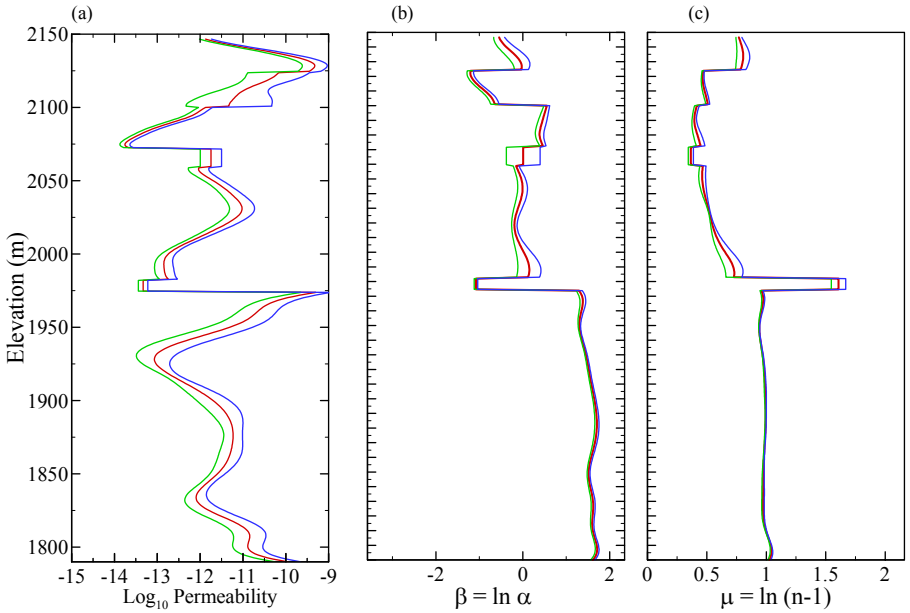

Figure 23. Estimated statistics of hydrologic properties: mean (red curves) and plus/minus two standard deviations (blue and green curves) for well LADP-4.

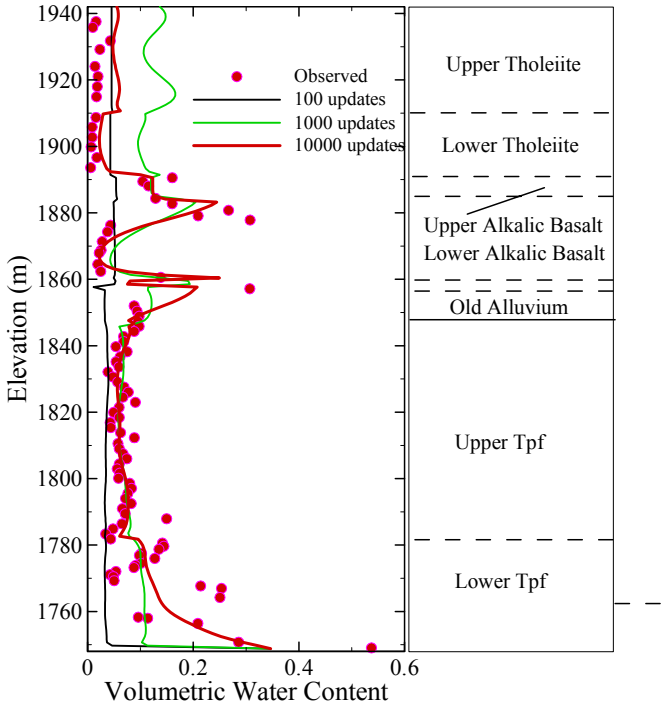

Figure 24. Comparison of observed water content with fitted profiles of different numbers of updates for well R-9.

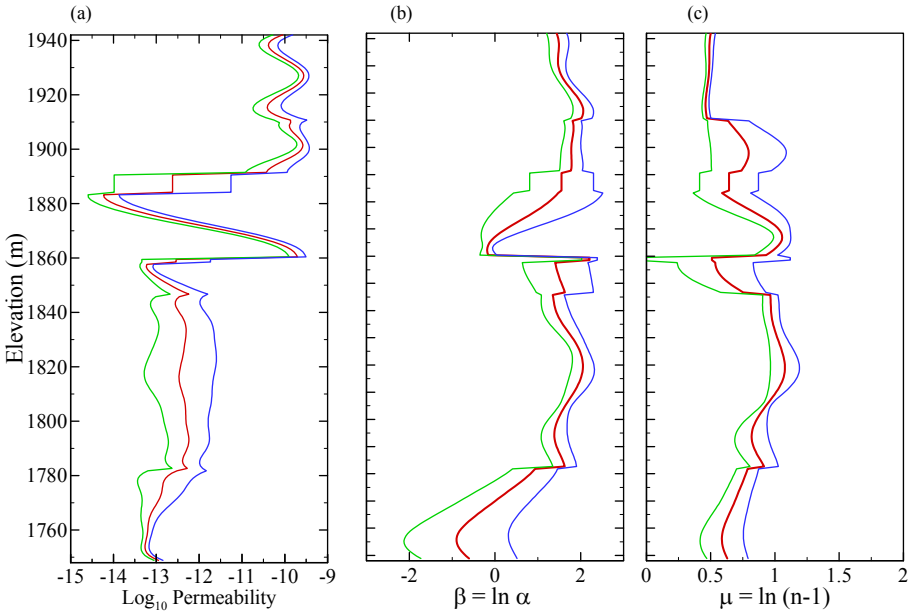

Figure 25. Estimated statistics of hydrologic properties: mean (red curves) and plus/minus two standard deviations (blue and green curves) for well R-9.

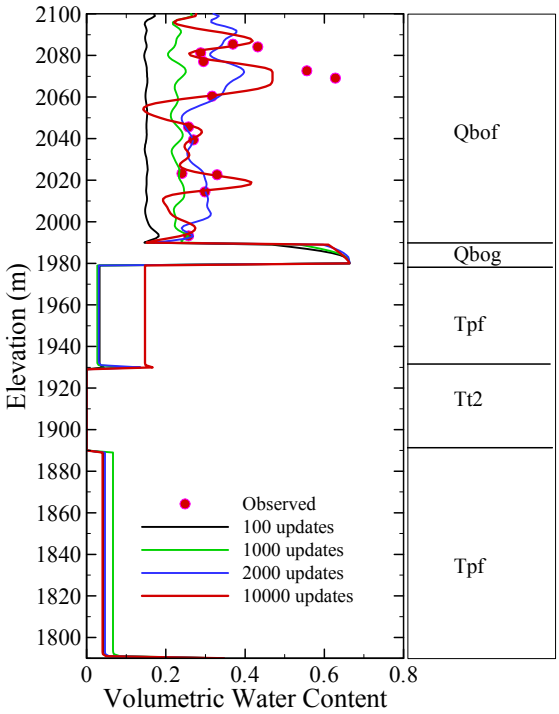

Figure 26. Comparison of observed water content with fitted profiles of different numbers of updates for well LAOI(A)-1.1.

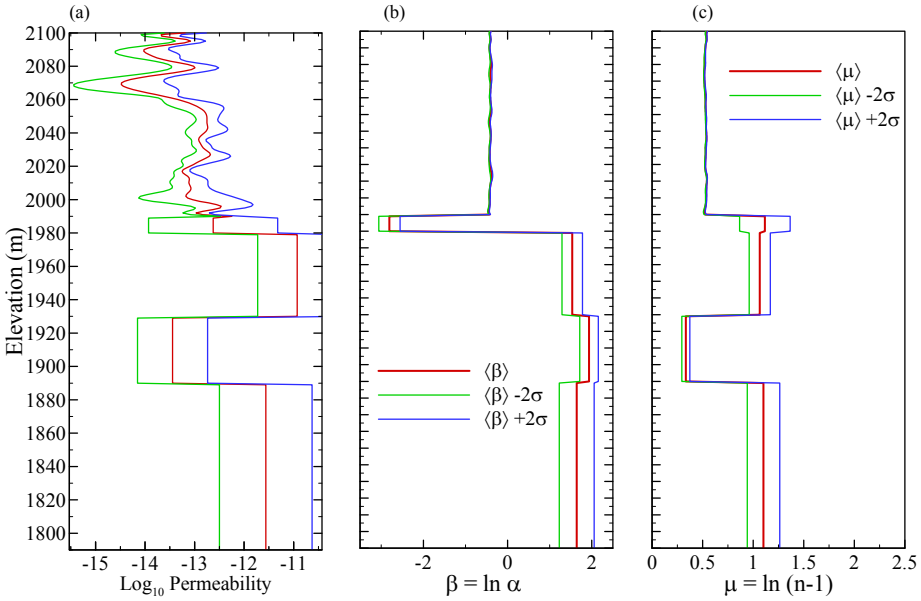

Figure 27. Estimated statistics of hydrologic properties: mean (red curves) and plus/minus two standard deviations (blue and green curves) for well LAOI(A)-1.1.

Table 3. Summary of Estimated Infiltration Statistics at Different Wells

Wells	Locations	Mean infiltration rate (mm/yr)	Std deviation of infiltration rate(mm/yr)	Comments
MCOBT-4.4	MC Canyon	125.3	30.9	
MCOBT-8.5	MC Canyon	21.9	5.7	No perched water
LADP-3	LA Canyon	220.0	16.7	
LADP-4	DP Canyon	10.9	1.9	No perched water
R-9	LA Canyon	139.9	84.2	
LAOI(A)-1.1	LA Canyon	523.7	45.9	

III. Discussion and Conclusions

The summary of infiltration rates presented in Table 3 illustrates the wide range of infiltration values obtained from water content profiles, depending on the topographic setting and location within canyons. Overall, the results are in general agreement with past analyses of infiltration rates at these and other similar locations on the Pajarito Plateau. For example, the difference in infiltration rate estimated between LADP-3 and LADP-4 illustrates the vast difference in downward percolation flux depending on whether the location in a wet canyon or a mesa/dry canyon. Within Los Alamos Canyon itself, the difference between LADP-3 and LAOI(A)-1.1 is thought to be due to the proximity of the latter well to a more intensely fractured region associated with the Guaje Mountain fault zone (Gray, 1997). The estimated value in R-9, located further down canyon, is lower than that of either of these two wells. However, the infiltration into highly fractured basalts with low permeability matrix may violate the basic assumption of porous flow in some of the rock units, making this estimate potentially suspect.

In Mortandad Canyon, the relative infiltration rates in the two wells are in keeping with the locations of these wells. MCOBT-4.4 is located further up canyon, and it may be that the source of water in MCOBT-8.5 may be depleted by infiltration and ET processes at this down-canyon location. Another factor to consider in Mortandad Canyon is that the analyses in these intermediate wells are focused on water content values in the upper part of the stratigraphic section. Recent reductions in the water discharged from the Radioactive Liquid Waste Treatment Facility (RLWTF) may have already impacted the water content values in the upper parts of the vadose zone where the measurements are made. If this is the case, then the infiltration rate estimates reflect the present-day infiltration rate, rather than the historical, presumably higher, infiltration rate. This concept of declining infiltration rate over time due to changes in the operation of the RLWTF has been promoted by Kwicklis et al. (2005), who applied the concept to an interpretation of several tritium peaks in well R-15. In principle, a transient analysis of the water content information could be performed for these wells to examine this possibility, but this approach was beyond the scope of this study. The analysis developed herein should be extended to this and other wells in Mortandad Canyon to evaluate the water content and contaminant profile information.

With regard to the new analysis technique, the MCMC method proved to be a very effective means for determining the mean and standard deviation of infiltration from the water content profiles. Obtaining the standard deviation is an important advance because many inverse techniques yield unrealistic estimates of the uncertainty. The MCMC method should provide appropriate uncertainty estimates because it takes into account heterogeneity in the medium properties, using that as a basis for obtaining better fits to the water content data. Numerical experiments with the method illustrate that including small-scale heterogeneity is important for improved matches to the data. When applied to the data from wells in the vicinity of LANL, the method provided reasonable estimates of uncertainty for wells in a variety of topographic settings. The more advanced versions of the method that include transient flow and solute transport should be useful for interpreting data currently being collected as part of the ER activities in Mortandad Canyon and other locations around the Laboratory.

IV. References

Birdsell, K. H., A. V. Wolfsberg, D. Hollis, T. A. Cherry, and K.M. Bower, 2000. Groundwater flow and radionuclide transport calculations for a performance assessment of a low-level waste site," *J. Contam. Hydrol.*, 46, 99-129.

Birdsell, K. H., B. D. Newman, D. E. Broxton, and B. A. Robinson, 2005, this issue. Conceptual models of vadose-zone flow and transport beneath the Pajarito Plateau, Los Alamos, New Mexico. *Vadose Zone J.*.

Birdsell, K. H., T. A. Cherry, P. Lichtner, and B. J. Travis, September, 1999. Numerical model of flow and transport for Area 2, MDA AB at TA-49," Los Alamos National Laboratory Report LA-UR-99-5501, Los Alamos, New Mexico.

Broxton, D. D. Vaniman, P. Longmire, B. Newman, W. Stone, A. Crowder, P. Schuh, R. Lawrence, E. Tow, M. Everett, R. Warren, N. Clayton, D. Counce, E. Kluk, and D. Bergfeld, 2002, Characterization Well MCOBT-4.4 and Borehole MCOBT8.5 Completion Report, Los Alamos National Laboratory report LA-13993-MS, Los Alamos, New Mexico.

Broxton, D., R. Gilkeson, P. Longmire, J. Marin, R. Warren, D. Naniman, A. Crowder, B. Newman, B. Lowry, D. Rogers, W. Stone, S. McLin, G. WoldeGabriel, D. Daymon, and D. Wycoff, 2001, Characterization Well R-9 Completion Report, Los Alamos National Laboratory report LA-13742-MS, Los Alamos, New Mexico.

Broxton, D. E., and D. T. Vaniman, 2005, this issue. Geologic framework of a groundwater system on the margin of a rift basin, Pajarito Plateau, North-Central New Mexico, *Vadose Zone J.*

Broxton, D.E., Longmire, P.A., Eller, P.G., and Flores, D., 1995, Preliminary drilling results for boreholes LADP-3 and LADP-4 *in* D.E.Broxton and P.G. Eller, eds., Earth Science Investigations for Environmental Restoration—Los Alamos National Laboratory Technical Area 21, Los Alamos National Laboratory Report LA-12934-MS pp. 93-110.

Carrera, J., and S. P. Neuman, Estimation of aquifer parameters under transient and steady state conditions, 1. Maximum likelihood method incorporating prior information, Water Resour. Res.,22(2), 199-210, 1986.

Clifton, P. M., and S. P. Neuman, Effects of kriging and inverse modeling on conditional simulation of the Avra Valley aquifer in southern Arizona, Water Resour. Res., 18(4), 1215-1234, 1982.

Dagan, G., and Y. Rubin, Stochastic identification of recharge, transmissivity, and storativity in aquifer transient flow: A quasi-state approach, Water Resour. Res., 24(10), 1698-1710, 1988.

Dagan, G., Stochastic modeling of groundwater flow by unconditional and conditional probabilities: The inverse problem, Water Resour. Res., 21(1), 65-72, 1985.

Dander, D. C., 1998. Unsaturated groundwater flow beneath upper Mortandad Canyon, Los Alamos, New Mexico, Los Alamos National Laboratory report LA-UR-98-4759, Los Alamos, New Mexico.

Ginn, T. R., J. H. Chushman, Inverse method for subsurfacr flow: a critical review of stochastic techniques, Stoch. Hydrol. HYdraul., 4, 1-26.

Gomez-Hernanez, J. J., A. S. Sahuquillo, and J. E. Capilla, Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data--I. Theory, J. Hydrology, 203, 162-174, 1997.

Gray, R. N., 1997. Hydrologic budget analysis and numerical simulations of groundwater flow in Los Alamos canyon near Los Alamos, New Mexico, Master's Thesis, University of New Mexico, Albuquerque, New Mexico.

Gutjahr, A., and J. Wilson, Co-kriging for stochastic flow models, Transport in Porous Media, 4, 585-598.

Hanna, S., and T-CJ. Yeh, Estimation of co-conditional moments of transmissivity, hydraulic head, and velocity fields, Adv. Water Resour., 22(1), 87-95, 1998.

Harvey C. F., and S. M. Gorelick, Mapping hydraulic conductivity: Sequential conditioning with measurements of solute arrival time, hydraulic head, and local conductivity, Water Resour. Res., 31(7), 1615-1626, 1995.

Hoeksema, R. J., and P. K. Kitanidis, An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling, Water Resour. Res., 20(7), 1003-1020, 1984.

Hoeksema, R. J., and P. K. Kitanidis, Comparison of Gaussian conditional mean and kriging estimation in the geostatistical solution of the inverse problem, Water Resour. Res., 21(6), 825-836, 1985.

Hughson, D. L., and A. Gutjahr, Effect of conditioning randomly heterogeneous transmissivity on temporal hydraulic head measurements in transient two-dimensional aquifer flow, Stoch. Hydrol. and Hydraul., 12, 155-170, 1998.

Kitanidis, P. K., and E. G. Vomvoris, A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations, Water Resour. Res., 19(3), 677-690, 1983.

Kitanidis, P. K., On the geostatistical approach to the inverse problem, Adv. Water Resour. 19(6), 333-342, 1996.

Kitanidis, P. K., Quasi-linear geostatistical theory for inversing, Water Resour. Res., 31(10), 2411-2419, 1995.

Kwicklis, E., M. Witkowski, K. Birdsell, B. Newman, and D. Walther, 2005. Development of an infiltration map for the Los Alamos area, New Mexico, Vadose Zone J., 4, 672-693.

LaVenne, A. M., B. S. RamaRao, G. deMarsily, M. G. Marietta, Pilot point methodology to automated calibration of an ensemble of conditionally simulated transmissivity fields, 2. Application, Water Resour. Res., 31(3), 495-516, 1995.

McLaughlin, D., and L. R. Townley, A reassessment of the groundwater inverse problem, Water Resour. Res., 32(5), 1131-1162, 1996.

Neuman, S. P., A statistical approach to the inverse problem of aquifer hydrology, 3, Improved solution method and added perspective, Water Resour. Res., 16(2), 331-346, 1980.

Oliver, D. S., L. B. Cunha, and A. C. Reynolds, Markov chain Monte Carlo methods for conditioning permeability field to pressure data, Math. Geol., 29(1), 61-91, 1997.

Rogers, D. B., and B. M.Gallaher, 1995. The Unsaturated Hydraulic Characteristics of the Bandelier Tuff, Los Alamos National Laboratory report LA-12968-MS, Los Alamos, New Mexico.

Rogers, D.B., B.M. Gallaher and E.L. Vold, 1996. Vadose zone infiltration beneath the Pajarito Plateau at Los Alamos National Laboratory," New Mexico Geological Society Guidebook, 47th Field Conference, Jemez Mountains Region, pp. 413-420.

Sun, N.-Z., Inverse Problems in Groundwater Modeling, Kluwer, Dordrecht, 1994.

van Genuchten, M. T., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Amer. J.*, 44 892- 898.

Zyvoloski, G.A., B.A. Robinson, Z.V. Dash, and L.L. Trease, 1997. Summary of the models and methods for the FEHM application - A finite-element heat- and mass-transfer code. Report LA-13307-MS, Los Alamos National Laboratory, Los Alamos, New Mexico.